

Improved velocimetry in optical coherence tomography using Bayesian analysis

Kevin C. Zhou,¹ Brendan K. Huang,¹ Hemant Tagare,^{1,2,3} and Michael A. Choma^{1,2,4,5,*}

¹Department of Biomedical Engineering, Yale University, 55 Prospect St., New Haven, Connecticut 06520, USA

²Department of Diagnostic Radiology, Yale University, 333 Cedar St., New Haven, Connecticut 06510, USA

³Department of Electrical Engineering, Yale University, 10 Hillhouse Ave., New Haven, Connecticut 06520, USA

⁴Department of Pediatrics, Yale University, 333 Cedar St., New Haven, Connecticut 06510, USA

⁵Department of Applied Physics, Yale University, P.O. Box 208267, New Haven, Connecticut 06520, USA

*michael.choma@yale.edu

Abstract: OCT is a popular cross-sectional microscale imaging modality in medicine and biology. While structural imaging using OCT is a mature technology in many respects, flow and motion estimation using OCT remains an intense area of research. In particular, there is keen interest in maximizing information extraction from the complex-valued OCT signal. Here, we introduce a Bayesian framework into the data workflow in OCT-based velocimetry. We demonstrate that using prior information in this Bayesian framework can significantly improve velocity estimate precision in a correlation-based, model-based framework for Doppler and transverse velocimetry. We show results in calibrated flow phantoms as well as *in vivo* in a *Drosophila melanogaster* (fruit fly) heart. Thus, our work improves upon the current approaches in terms of improved information extraction from the complex-valued OCT signal.

©2015 Optical Society of America

OCIS codes: (110.4500) Optical coherence tomography; (000.5490) Probability theory, stochastic processes, and statistics; (110.4153) Motion estimation and optical flow; (170.3880) Medical and biological imaging; (290.5820) Scattering measurements; (030.6140) Speckle.

References and links

1. W. Drexler, M. Liu, A. Kumar, T. Kamali, A. Unterhuber, and R. A. Leitgeb, "Optical coherence tomography today: speed, contrast, and multimodality," *J. Biomed. Opt.* **19**(7), 071412 (2014).
2. B. J. Berne and R. Pecora, *Dynamic Light Scattering* (John Wiley & Sons, Inc., New York, 1976).
3. Y. Imai and K. Tanaka, "Direct velocity sensing of flow distribution based on low-coherence interferometry," *J. Opt. Soc. Am. A* **16**(8), 2007–2012 (1999).
4. J. Lee, W. Wu, J. Y. Jiang, B. Zhu, and D. A. Boas, "Dynamic light scattering optical coherence tomography," *Opt. Express* **20**(20), 22262–22277 (2012).
5. V. J. Srinivasan, H. Radhakrishnan, E. H. Lo, E. T. Mandeville, J. Y. Jiang, S. Barry, and A. E. Cable, "OCT methods for capillary velocimetry," *Biomed. Opt. Express* **3**(3), 612–629 (2012).
6. X. Liu, Y. Huang, J. C. Ramella-Roman, S. A. Mathews, and J. U. Kang, "Quantitative transverse flow measurement using optical coherence tomography speckle decorrelation analysis," *Opt. Lett.* **38**(5), 805–807 (2013).
7. N. Weiss, T. G. van Leeuwen, and J. Kalkman, "Localized measurement of longitudinal and transverse flow velocities in colloidal suspensions using optical coherence tomography," *Phys. Rev. E* **88**(4), 042312 (2013).
8. B. K. Huang and M. A. Choma, "Resolving directional ambiguity in dynamic light scattering-based transverse motion velocimetry in optical coherence tomography," *Opt. Lett.* **39**(3), 521–524 (2014).
9. B. K. Huang, U. A. Gamm, V. Bhandari, M. K. Khokha, and M. A. Choma, "Three-dimensional, three-vector-component velocimetry of cilia-driven fluid flow using correlation-based approaches in optical coherence tomography," *Biomed. Opt. Express* **6**(9), 3515–3538 (2015).
10. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.* **21**(6), 1087–1092 (1953).
11. M. Plummer, "JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling" in *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, F. K. Hornik, ed. (Technische Universität Wien, Vienna, Austria, 2003).
12. M. Steyvers, "MATJAGS: a Matlab interface for JAGS," (2011), http://psiexp.ss.uci.edu/research/programs_data/jags/.

13. A. Charnes, E. L. Frome, and P. L. Yu, "The Equivalence of Generalized Least Squares and Maximum Likelihood Estimates in the Exponential Family," *J. Am. Stat. Assoc.* **71**(353), 169–171 (1976).
14. C. Kasai, K. Namekawa, A. Koyano, and R. Omoto, "Real-Time Two-Dimensional Blood Flow Imaging Using an Autocorrelation Technique," *IEEE Trans. Sonics and Ultrasonics* **32**(3), 458–464 (1985).
15. V. X. D. Yang, M. L. Gordon, A. Mok, Y. Zhao, Z. Chen, R. S. C. Cobbold, B. C. Wilson, and I. Alex Vitkin, "Improved phase-resolved optical Doppler tomography using the Kasai velocity estimator and histogram segmentation," *Opt. Commun.* **208**(4-6), 209–214 (2002).
16. S. Geman and C. Graffigne, "Markov Random Field Image Models and Their Applications to Computer Vision," in *International Congress of Mathematicians*, A. M. Gleason, ed. (American Mathematical Society, Berkeley, California, 1986), pp. 1496–1517.

1. Introduction

Optical coherence tomography (OCT) is widely used for flow velocity estimation in biomedical applications [1]. Doppler-based approaches are a popular technique used to estimate velocity. However, traditional Doppler-OCT is only able to quantify axial flow velocities given that the Doppler signal is proportional to the dot product of the velocity vector and a unit vector along the optical axis. Several newer correlation-based approaches exploit models of the complex-valued OCT signal that incorporate physical models of scatterer motion into models of image formation. Modeling the complex-valued, time-varying correlation signal enables the extraction of flow velocity information orthogonal to the optical axis, including total flow speed [2–7] and directional velocimetry along three spatial axes [8, 9].

As with most approaches to velocimetry, both Doppler and correlation-based approaches require multiple measurements taken in time in order to generate speed or velocity estimates. The need for multiple measurements increases total imaging time. Approaches that improve measurement precision while minimizing the need for repeated measures are therefore of keen interest in OCT-based velocimetry. Here, using an autocorrelation model of the time-varying OCT signal, we perform axial and transverse velocimetry using a computational Bayesian approach known as Markov chain Monte Carlo (MCMC) [10]. The Bayesian MCMC approach generates flow velocity estimates along with their associated uncertainties. We further show that our model-based approach allows natural ways of incorporating prior knowledge of system and sample parameters, thereby improving estimation precision. Thus, the incorporation of prior knowledge using a Bayesian framework enables improved velocity estimation precision compared to prior non-parametric approaches to Doppler and correlation-based OCT velocimetry. These prior approaches are non-parametric in the sense that they do not assume an underlying noise model in its estimation procedure. Likewise, a Bayesian framework can enable fewer repeated measures (less data) without sacrificing velocity estimation precision.

2. Brief description of the complex-valued OCT signal, the complex-valued autocorrelation signal, and directional dynamic light scattering OCT (DLS-OCT)

For a scattering sample consisting of M scatterers, the time-varying complex-valued OCT signal at location \mathbf{r}_0 and time $t \geq 0$ can be modeled as a complex phasor summation:

$$i(\mathbf{r}, t) = \sum_{m=1}^M \int \text{psf}(\mathbf{r}') \delta(\mathbf{r}' - [\mathbf{r} - \mathbf{r}_m(t)]) d\mathbf{r}' \quad (1a)$$

$$\mathbf{r}_m(t) = \mathbf{r}_m(0) + \mathbf{v}t \quad (1b)$$

\mathbf{r} is the Cartesian coordinate vector (x, y, z), \mathbf{r}' is a dummy variable used to exploit the sifting property of the Dirac delta function $\delta(\mathbf{r})$, $\text{psf}(\mathbf{r})$ is the point spread function of the imaging system, and \mathbf{v} is the temporally stationary vectorial velocity of the imaged scatterers. $\text{psf}(\mathbf{r})$ is complex-valued and typically is modeled as a real-valued Gaussian amplitude envelope

modulated by a complex-valued fringe (Fig. 1). The fringe has the form $\exp(2jk[z_0 - z_m(t)])$, where $t \geq 0$, $j = (-1)^{1/2}$, and k is the optical wavenumber in radians per unit distance.

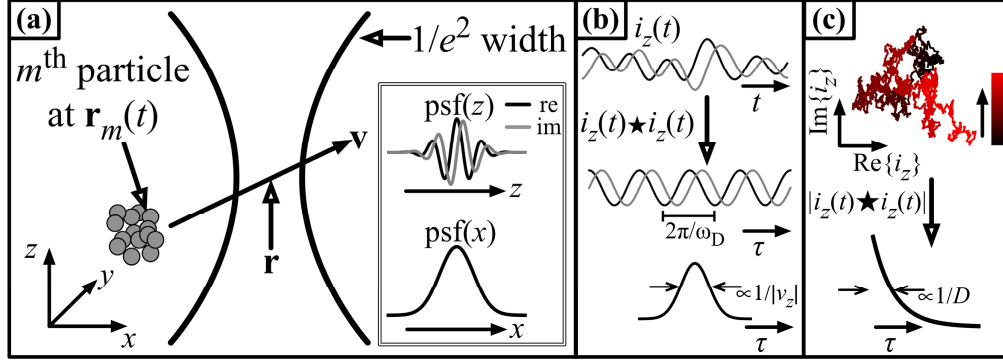


Fig. 1. (a) Model of the time-varying, complex-valued OCT signal $i(\mathbf{r})$. Each particle in an ensemble of M identical, uniformly moving, randomly distributed particles contributes to $i(\mathbf{r})$. The contribution is weighted by the point-spread function (psf), which is complex-valued in the z -axis and real-valued in the x - and y -axes. Typically, $\text{psf}(y) = \text{psf}(x)$. For clarity, only a subset of the M scatterers are shown. The scatterers also undergo diffusion with a root mean squared displacement given by a diffusivity parameter D . (b) $i_z(t)$ is the axial response, that is, the complex-valued OCT signal along the z -axis. As shown in Eq. (2), autocorrelation of the complex-valued signal yields a complex-valued fringe as well as an amplitude envelope. Here, the \star operator indicates correlation. The frequency of the fringe is given by the Doppler shift imparted by the axial component of the moving scatterers. The amplitude envelope reproduces the shape of the point-spread function envelope. The width of the envelope is modulated by the axial speed (v_z), that is, the magnitude of the axial velocity. Assuming $\text{psf}(y) = \text{psf}(x)$, the real-valued response along the x - and y -axes is likewise an amplitude envelope with a width modulated by the total in-plane speed $(v_x^2 + v_y^2)^{1/2}$. (c) In the case of purely diffusive motion of monodisperse scatterers in the axial direction, the magnitude of the autocorrelation of the complex-valued signal is an exponential decay. The characteristic decay time is inversely proportional to the particle diffusivity. For short periods of time (shown here), the signal may wander in a local neighborhood. Over time, the signal fills out speckle statistics in the complex plane.

The evolution of $i(\mathbf{r}, t)$ in time is a stochastic process in the complex plane. If, however, the spatial distribution of particles is a white noise process, a functional form of the autocorrelation of $i(\mathbf{r}, t)$ can be written. If we assume that particle motion has a diffusive component and a linear translational component, the autocorrelation of $i(\mathbf{r}, t)$ can be modeled as [4]:

$$G(\mathbf{r}, \tau) = \text{Re} \exp(2jkv_z\tau) \exp\left(-\frac{v_x^2 + v_y^2\tau^2}{w_{xy}^2} - \frac{v_z^2\tau^2}{w_z^2}\right) \exp(-4k^2D|\tau|) \quad (2)$$

τ is the autocorrelation lag, D is particle diffusivity, w_{xy} is the $1/e^2$ beam radius in the x - y plane, w_z is the $1/e^2$ longitudinal coherence length, and \mathbf{v} is the vector component of \mathbf{v} in a given Cartesian direction. We assume a Gaussian form of the axial and transverse point spread functions. Every parameter is an implied function of \mathbf{r} . Several estimators (e.g. Kasai) focus on the $\exp(2jkv_z\tau)$ term since $2kv_z$ represents the Doppler shift in units of radians per second. These estimators are limited to axial velocity estimation. By fitting the numerical autocorrelation of the complex-valued OCT signal to this model, the total speed $(v_x^2 + v_y^2 + v_z^2)^{1/2}$ of the flow can be determined as well as the axial velocity v_z . This approach is called dynamic light scattering OCT (DLS-OCT) [4]. Further, if three frames of reference are recognized (i.e. sample [*samp*], scanner [*scan*], and detector [*det*]), if Eq. (2) is written in terms of $v_{x,\text{det}} = v_{x,\text{scan}} - v_{x,\text{samp}}$, and if the signal is acquired at a series of different scan rates

$v_{x,scan}$, $v_{x,samp}$ can be estimated (Fig. 2). We term this approach directional DLS-OCT. Hence, the model can be rewritten as [8]:

$$G(\mathbf{r}, v_x^{scan}, \tau) = R \exp(2jkv_2\tau) \exp\left(-\frac{\left((v_x - v_{x,scan})^2 + C\right)\tau^2}{w_{xy}^2}\right) \quad (3)$$

C is the baseline decorrelation rate resulting from velocity components orthogonal to the scan bias direction and from diffusion. Note that we simplified our model by lumping diffusion together with velocity-based decorrelation, even though diffusion contributes as a single exponential decorrelation. Nevertheless, the interpretation of Eq. (3) is that as $v_{x,scan}$ is varied, the magnitude of the decorrelation rate is modulated (Fig. 2). The magnitude of decorrelation is minimized when flow velocity matches scan velocity in the x -direction, yielding an estimate of $v_{x,samp}$. While Eq. (3) does not explicitly recover an estimate of scatterer diffusivity, Eq. (3) in conjunction with a scan bias modulation protocol affords a mechanism for separating a single transverse velocity component (i.e. v_x) from all other sources of translational and diffusive motion. Because our model acknowledges both translational and diffusive scatterer motion, we chose to use the Lee, *et al.* nomenclature of DLS-OCT [4].

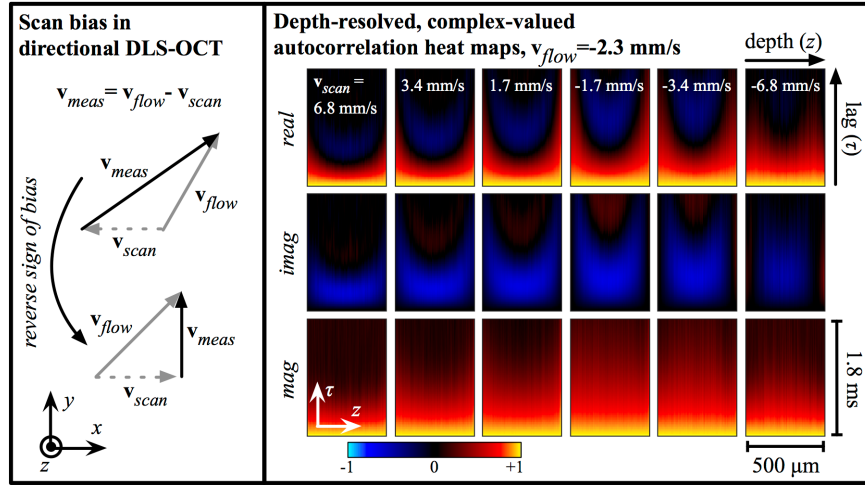


Fig. 2. Left panel: The measured velocity is the difference between flow velocity in the object being imaged (\mathbf{v}_{flow}) and the velocity of the beam scanner that deflects the imaging beam (\mathbf{v}_{scan}). Varying \mathbf{v}_{scan} breaks a symmetry that is otherwise present in DLS-OCT. Symmetry is broken because \mathbf{v}_{meas} is different when sign of \mathbf{v}_{scan} is flipped. Moreover, the velocity component along the x -axis (or y -axis) can be estimated by exploiting the fact that \mathbf{v}_{meas} is minimized when $\mathbf{v}_{flow} = \mathbf{v}_{scan}$. Right panel: Value of the autocorrelation of the complex-valued OCT signal as a function of time lag (τ) and axial location (z). Data is from a calibrated flow phantom described in Section 4.1. Here, the direction of \mathbf{v}_{scan} defines the x -axis, and \mathbf{v}_{flow} is nominally parallel to the x -axis. Decorrelation times are longer (i.e. rates of decorrelation are slower) as \mathbf{v}_{scan} approaches \mathbf{v}_{flow} . Note that the fringe frequency (Doppler shift) varies slowly with scan bias speed, indicating a small beam scanning-induced Doppler shift. By analyzing the total Doppler shift as a function of scan bias velocity, we estimate that the scanner-induced Doppler shift equivalent to $-16.4 \mu\text{m/s}$ per 1 mm/s of scan velocity.

3. Bayesian framework, modeling, and analysis

3.1 General framework and noise model

The goal of OCT velocimetry is to generate spatially indexed velocity maps. Bayesian analysis can improve upon existing approaches to OCT velocimetry by (a) providing probability density functions that represents the uncertainty in velocity estimates and (b)

giving a framework for the incorporation of prior information. One expression of Bayes' rule states that

$$P(\theta|data) = \frac{P(data|\theta)P(\theta)}{P(data)} \quad (4a)$$

$$P(data) = \int P(data|\theta)P(\theta)d\theta \quad (4b)$$

Here, θ is a vector of the estimated parameters, $P(\theta|data)$ is the posterior distribution, $P(data|\theta)$ is the likelihood function, $P(\theta)$ is the prior distribution, and $P(data)$ is a normalization factor that ensures that the posterior distribution integrates to unity. The prior distribution represents prior knowledge about the probability of parameter values in the absence of data. The likelihood function is the statistical model of noisy data. It models how the noisy data is dependent on the parameters. Moreover, in the context of maximum likelihood estimation, $P(data|\theta)$ viewed as a function of θ is called the likelihood function

$L(\theta)$ that is used to estimate $\theta_{MLE} = \arg \max_{\theta} L(\theta)$. The posterior distribution represents an

estimate of the parameters given the data. The dimensionality of the posterior distribution is equal to the number of parameters in θ . While the full posterior distribution can be difficult to visualize, it is useful to obtain the one-dimensional function $P(\theta_n|data)$, the marginal distribution, that represents the probability density function for a single parameter θ_n given the data:

$$P(\theta_n|data) = \int_{\theta_n^c} P(\theta|data)d\theta_n^c \quad (5a)$$

$$\theta = \theta_n \quad \theta_n^c \quad (5b)$$

Here, the posterior distribution is marginalized over all parameters except θ_n . The widths of the one-dimensional probability density function $P(\theta_n|data)$ is an intuitive measure of the uncertainty θ_n .

In our approach we estimate the posterior probability in a pointwise manner. That is, we estimate $P_{r_0}(data|\theta)$ at individual locations r_0 . We assume a homoscedastic Gaussian noise model for the likelihood function $P_{r_0}(data|\theta)$. Specifically,

$$P_{r_0}(data|\theta) = \prod_{\tau, v_{x,scan}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left|G_{r_0}(\tau, v_{x,scan}) - data_{r_0}(\tau, v_{x,scan})\right|^2}{2\sigma^2}\right) \quad (6a)$$

$$\theta = \{R, v_x, w_{xy}, C, \sigma^2\} \quad (6b)$$

Here, σ^2 is the variance of the noise in the data.

Note that the exponent of $P_{r_0}(data|\theta)$ is a function of the model of the complex autocorrelation function G (Eqs. (2) or (3)). As a consequence, regardless of the simplicity of the form of the prior distribution $P(\theta)$, evaluating the integral in Eq. (4b) to estimate $P(data)$ is a non-trivial task. As such, we used JAGS (Just Another Gibbs Sampler [11]) in MATLAB (MATJAGS [12]) to numerically integrate Eq. (4b) and estimate $P_{r_0}(\theta|data)$. We used a Markov chain Monte Carlo (MCMC) approach [10] in MATJAGS for numerical integration. Each MCMC run was of length 10^4 , a histogram of the parameters of which gives the

estimate of the posterior distribution. MCMC is a numerical integration technique that commonly is used to obtain posterior distributions in numerical Bayesian analysis. Even though a mathematical expression for the posterior distribution can be defined, the denominator of that expression (i.e. $P(data)$), is in general very difficult to directly calculate. MCMC implicitly estimates $P(data)$ using a numerical approach. MCMC executes a random walk across the parameter space in proportion to the posterior density. In this way, the Markov chain is drawn towards regions in the parameter space with high posterior probability and visits the lower probability regions proportionally less often. Hence, the number of times that the Markov chain visits each location in the parameter space gives an estimate of the posterior probability. In our experience, MCMC run lengths of 10^3 typically reach steady state. We used run lengths of 10^4 to ensure that not reaching steady state is a remote possibility.

3.2 Incorporation of prior information through the use of an adaptive hyperprior

In laminar flow as well as with other well-behaved motions, a continuity argument suggests that the parameter values of neighboring locations are similar. One way to formally incorporate such a continuity argument into Bayes' rule (Eq. (4)) is through an adaptive hyperprior. If $P(\theta)$ is a Gaussian distribution, then a hyperprior defines the Gaussian distribution not in terms of fixed values for its mean and variance but rather the mean and variances are themselves taken as random variables from another distribution called the hyperprior distribution. In our case the hyperprior itself is a Gaussian distribution. At current (*curr*) location \mathbf{r}_o , we use the posterior distribution at a neighboring (*neigh*) location as a hyperprior on the prior distribution at \mathbf{r}_o . Using a hyperprior, then, Bayes' rule (Eqs. (4a) and (4b)) expands to:

$$P(\theta_{curr}, \theta_{neigh} | data_{curr}) = \frac{P(data_{curr} | \theta_{curr}) P(\theta_{curr} | \theta_{neigh}) P(\theta_{neigh})}{P(data_{curr})} \quad (7a)$$

$$P(data_{curr}) = \iint P(data_{curr} | \theta_{curr}) P(\theta_{curr} | \theta_{neigh}) P(\theta_{neigh}) d\theta_{curr} d\theta_{neigh} \quad (7b)$$

Likewise, the one-dimensional function $P(\theta_{n,curr} | data_{curr})$ that represents the probability density function for a single parameter $\theta_{n,curr}$ given the data is

$$P(\theta_{curr,n} | data_{curr}) = \iint P(data_{curr}, \theta_{curr} | data_{curr}) d\theta_{curr,n}^c d\theta_{neigh} \quad (8a)$$

$$\theta_{curr} = \theta_{curr,n} \quad \theta_{curr,n}^c \quad (8b)$$

$P(\theta_{curr} | \theta_{neigh})$ is a Gaussian distribution centered at θ_{neigh} , the value of which is governed by the hyperprior distribution $P(\theta_{neigh})$. The prior and hyperprior distributions have the same width, given by the width of the posterior of the neighbor. Hence, by progressively moving across an image, adapting the hyperprior on the mean of the prior at the current location to the posterior distribution at the neighboring location, we can reduce uncertainty in our velocity estimates.

Our use of the uninformative hyperprior approach to establish a performance baseline for the adaptive hyperprior approach is supported by the following argument. As discussed above, a maximum likelihood estimate (MLE) is the argmax with respect to θ of the likelihood function. In the case of an uninformative hyperprior, $P(\theta)$ is a very broad Gaussian distribution, meaning that the posterior distribution $P(\theta | data)$ is essentially determined by the likelihood function. Additionally, maximum likelihood estimation using a homoscedastic Gaussian noise model is equivalent to least squares regression fitting [13]. Thus, the uninformative hyperprior case reflects information used in two widely used estimation

processes (MLE and least squares), supporting its use as a baseline for evaluating performance of the adaptive hyperprior Bayesian approach.

4. OCT imaging and Kasai Doppler processing

4.1 OCT data collection

We used a $\lambda_0 = 1325$ nm spectral domain OCT system (Thorlabs Telesto) to image a rectangular (0.5 x 5 mm) flow channel containing an aqueous suspension of 100-nm diameter polystyrene beads with 0.1% Tween to prevent bead aggregation. Two vector component flow velocity estimates (v_x, v_z) were generated along one spatial dimension (z -axis). The peak total speed was estimated to be -2.3 mm/s based on the bulk flow rate of the syringe pump and the flow channel geometry. The scan direction was nominally parallel to the flow direction (the x -direction). The A-scan rate was 28 kHz. For all scan biases, scanning was performed over a fixed range of 250 μm ; as such, the faster scan velocities result in fewer data points. Wild-type fruit fly (*Drosophila melanogaster*) M-mode images were collected at a 28 kHz A-scan rate. One vector component heart wall velocity estimates (v_z) were generated along one spatial dimension (z -axis).

4.2 Doppler estimation using the Kasai autocorrelation algorithm

The Kasai method [14, 15] estimates the Doppler frequency through Taylor expansion and algebraic manipulations of a model of the autocorrelation signal (e.g. Eq. (2)). In doing so, the phase is estimated through the ratio of the real and imaginary components of a single-lag autocorrelation calculation. The Doppler frequency is then approximated by the change in phase after the first time lag:

$$\omega_{dop} = \frac{1}{\tau_{lag}} \tan^{-1} \frac{\text{Im}\{G(\mathbf{r}, \tau_{lag})\}}{\text{Re}\{G(\mathbf{r}, \tau_{lag})\}} \quad (9)$$

5. Results

5.1 Doppler axial velocimetry in a calibrated flow phantom

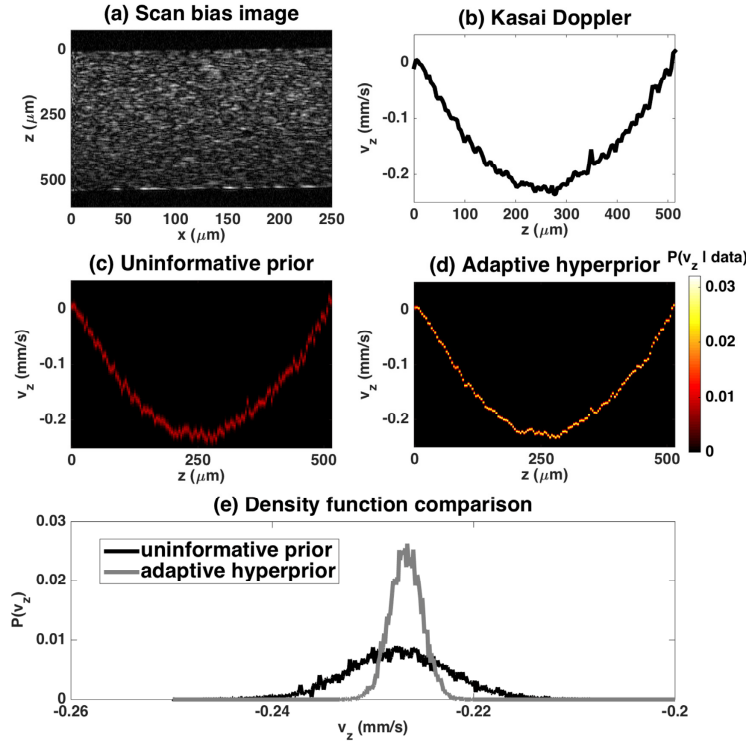


Fig. 3. Axial velocity (v_z) estimation using the time-varying signals acquired at a scan bias velocity of -1.7 mm/s. (a) Representative B-scan at a single scan bias velocity. (b) Kasai estimate of the axial velocity, (c) Bayesian analysis with an uninformative prior and (d) an adaptive hyperprior. (e) A sample uncertainty comparison at the center of the channel (blue = uninformative prior, green = adaptive hyperprior, red = Kasai). Here, the posterior distribution of the axial velocity $P(v_z | data)$ is defined in Eqs. (4) and (7).

The use of prior information through adaptive hyperpriors in a Bayesian framework improved Doppler-based velocity estimation (Fig. 3). We estimated flow velocity profiles using the standard Kasai Doppler estimator (Fig. 3(b)), using Bayesian analysis with a wide prior distribution (uninformative hyperprior; Fig. 3(c)), and using Bayesian analysis with adaptive hyperpriors (Fig. 3(d)). For our Bayesian analysis, Doppler frequencies were estimated by fitting the data to an equation that has the general form of Eq. (2) but that combines the last two terms on the right-hand side into a single Gaussian decay term. Since every possible axial velocity value at each lateral axial location has an associated probability (i.e. the posterior distribution $P(v_z | data)$), the Bayesian analysis results are displayed as heat maps. The width of the posterior distribution at each lateral axial location is an intuitive measure of the credibility of the velocity estimation process. These widths often are referred to as credible intervals (CIs). Comparing credible intervals demonstrates that the adaptive hyperprior significantly improves axial velocity estimation. If we define a reduction in uncertainty parameter

$$RU = \frac{CI_{95\%}^{UHP} - CI_{95\%}^{AHP}}{CI_{95\%}^{UHP}} \times 100\%, \quad (10)$$

there is a $RU = 70\%$ reduction in uncertainty attributable to the incorporation of prior information through an adaptive hyperprior. Here, $CI_{95\%}^{UHP}$ and $CI_{95\%}^{AHP}$ are the 95% CIs for using uninformative hyperpriors and adaptive hyperpriors, respectively. In the results in Fig. 3, $CI_{95\%}^{UHP} = 12\mu m$ and $CI_{95\%}^{AHP} = 6.3\mu m / s$.

As an additional comparison between the Bayesian estimates and Kasai Doppler estimate, we collapsed the posterior probability density functions to flow velocity profiles and compared them to Kasai Doppler profiles (Fig. 4). The posterior probability density functions were collapsed using a centroid calculation:

$$\bar{v}_z = \int v_z P(v_z | data) dv_z \quad (11)$$

The three Doppler flow velocity profiles are almost indistinguishable from each other (Fig. 4(a)). We additionally fit each flow velocity profile to a parabola. The three fits also are almost indistinguishable from each other (Fig. 4(b)). These results indicate that the posterior probability density functions contain information to generate flow profiles that are similar to those generated by a traditional Doppler estimator. We compared the collapsed posterior probability density functions to Kasai Doppler estimates because the Kasai estimator does not yield density functions that can be compared to full posterior probability density functions.

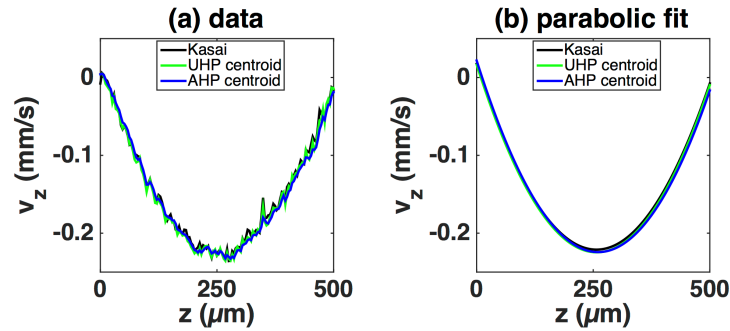


Fig. 4. Doppler flow velocity profiles generated using the Kasai Doppler estimator and by centroiding the posterior probability density functions for the uninformative hyperprior (UHP) and adaptive hyperprior (AHP) estimators. The R^2 values (minimum velocity values) for Kasai, UHP, and AHP parabolic fits are 0.989 (−0.221 mm/s), 0.991 (−0.225 mm/s), and 0.995 (−0.224 mm/s), respectively.

5.2 DLS-OCT total velocimetry in a calibrated flow phantom

The use of prior information through adaptive hyperpriors in a Bayesian framework also improved two-component flow velocity vector estimation in directional DLS-OCT. We investigated improvement in estimation performance in the context of varying the number of repeated data acquisitions (n_{data}) and varying the number of different scan bias velocities used in the directional DLS-OCT scan protocol (n_{bias}). Here, n_{data} indicates the number of images taken at each of n_{bias} scan biases used in the directional DLS-OCT scan protocol. We estimated the directional flow profile of a phantom calibrated to a peak flow of −2.3 mm/s with a near-90 degree Doppler angle. We used $n_{bias} = 8$ for fitting Eq. (3). to infer the flow velocity. Figure 2 shows the temporal autocorrelation of time-varying, complex-valued OCT signal across a calibrated flow phantom at different scan bias velocities.

From the computed autocorrelation functions at 8 scan biases ($n_{bias} = 8$), we then calculated the posterior distribution of the lateral flow velocity $P(v_x|data)$ and the axial flow velocity $P(v_z|data)$, based on Eqs. (4)–(8). We investigated whether an adaptive hyperprior could improve velocity estimation and the dependence of that improvement on n_{data} . Figure 5 shows the Bayesian estimates of v_x and v_z profiles along the z -axis (depth) using $n_{data} = 1$ and $n_{bias} = 8$ with uninformative and adaptive hyperpriors as well as using $n_{data} = 10$ and $n_{bias} = 8$.

with uninformative and adaptive hyperpriors. As with the Doppler results shown in Fig. 3, every possible axial location along the horizontal axis has an associated posterior probability density function of $P(v_x | \text{data})$ or $P(v_z | \text{data})$. As such, the data are displayed as heat maps. Overall, the magnitudes of the lateral and axial velocities correspond to a Doppler angle of 96° , in close agreement with geometric estimates (95°). For $n_{\text{data}} = 1$, there is a relatively high uncertainty in the reconstruction, but using an adaptive hyperprior reduces this uncertainty. For estimating v_x with $n_{\text{data}} = 1$, $CI_{95\%}^{\text{UHP}} = 1.81 \text{ mm/s}$ and $CI_{95\%}^{\text{AHP}} = 1.21 \text{ mm/s}$, giving $RU = 33\%$. Increasing the number of repeated data acquisitions significantly narrows credible intervals. For $n_{\text{data}} = 10$, $CI_{95\%}^{\text{UHP}} = 0.58 \text{ mm/s}$ and $CI_{95\%}^{\text{AHP}} = 0.40 \text{ mm/s}$, giving $RU = 31\%$. the 95% CIs were of widths $\sim 0.58 \text{ mm/s}$ and $\sim 0.40 \text{ mm/s}$. For axial velocity estimation with $n_{\text{data}} = 1$, $CI_{95\%}^{\text{UHP}} = 23 \mu\text{m/s}$ and $CI_{95\%}^{\text{AHP}} = 15 \mu\text{m/s}$, giving $RU = 35\%$. For axial velocity estimation with $n_{\text{data}} = 10$, $CI_{95\%}^{\text{UHP}} = 7.8 \mu\text{m/s}$ and $CI_{95\%}^{\text{AHP}} = 5.3 \mu\text{m/s}$, giving $RU = 32\%$. Thus, use of an adaptive hyperprior in DLS-OCT reduced uncertainties by approximately 30% in all of these cases.

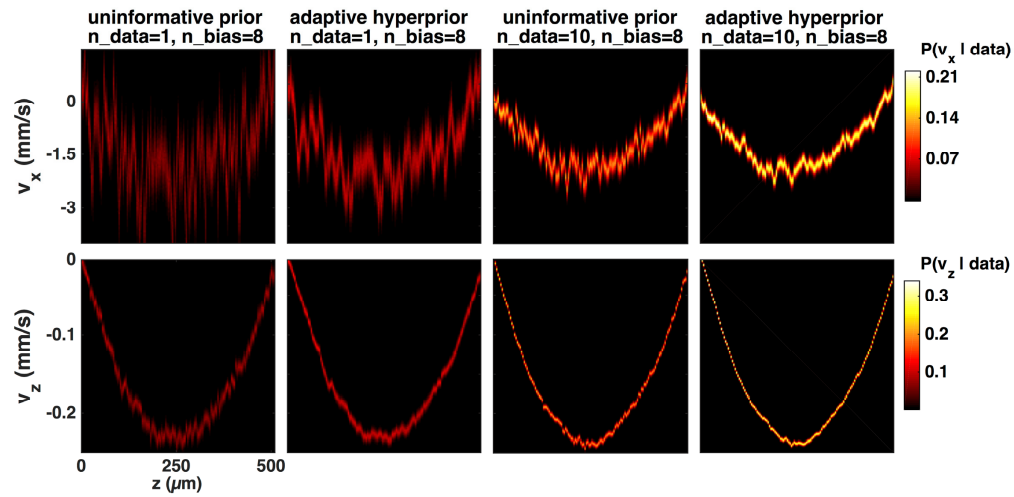


Fig. 5. Bayesian estimates of two-component flow velocity vectors: lateral flow velocity (v_x) and axial velocity (v_z). They were reconstructed using either an uninformative prior (i.e. very broad prior probability $P(\theta)$) or an adaptive hyperprior (i.e. prior probability is defined by a neighboring posterior distribution). Using a larger sample size and incorporating neighboring information improves the precision. We define precision by the width of the posterior probability density function. Each row of subfigures uses the same color bar. The posterior distribution of the lateral velocity $P(v_x | \text{data})$ and of the axial velocity $P(v_z | \text{data})$ is defined in Eqs. (4) and (7).

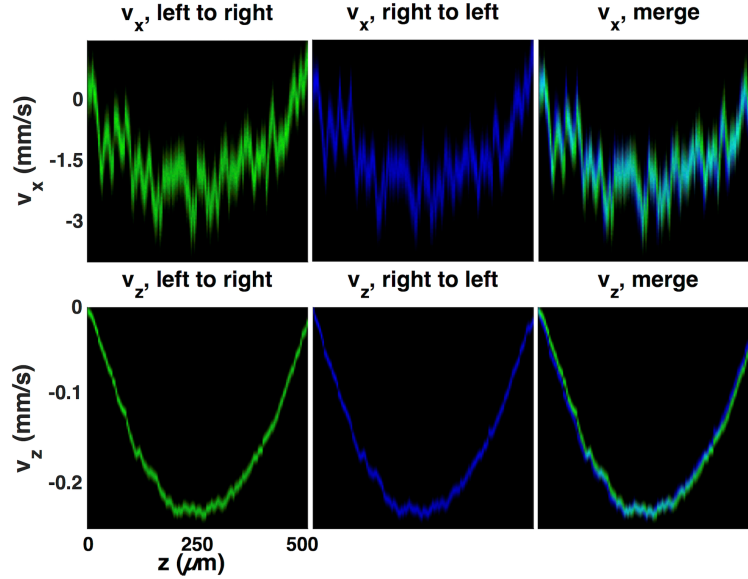


Fig. 6. Comparison of adaptive hyperprior estimation when the estimation process begins on the left-hand side and moves right (left to right; green colormap) and when it begins on the right-hand side and moves left (right to left; blue colormap). The flow velocity profiles are similar in either case. Data are from the $n_{data} = 1$, $n_{bias} = 8$ (second column in Fig. 5). The “merge” images are RGB color images in which the green channel is the left-to-right data and the blue channel is the right-to-left data. Cyan-appearing pixels in the “merge” images indicate a high degree of overlap between the left-to-right and right-to-left profiles. The left-to-right profile has a slight rightward shift and, likewise, the right-to-left profile has a slight leftward shift.

Our adaptive hyperprior estimation process moves left-to-right (i.e. from low values of z to high values of z). That is, when estimating velocity at a particular location, prior information is pulled from the adjacent pixel to the left. The first location in the estimation process uses an uninformative prior. In order to investigate the influence of moving left-to-right versus right-to-left during the estimation process, we generated velocity estimates in the $n_{data} = 1$, $n_{bias} = 8$ case in Fig. 5 when moving in each direction (Fig. 6). The v_x and v_z velocity profiles were similar in each case. The velocity profiles were slightly spatially shifted from each other, suggesting a small spatial lag in the estimation process similar to that observed with a low-pass filter (e.g. moving average filter).

Lastly, our beam waist estimates (derived from the autocorrelation signal as modeled in Eq. (3)) are consistent with but not equal to the imaging system resolution (as determined from intensity images of sparsely distributed sub-resolution scatterers). The $1/e^2$ beam radius values estimated using the autocorrelation approach described in this manuscript were in the 4.5 to 5 μm range. The $1/e^2$ beam radius as ascertained from imaging sub-resolution scatterers is $\sim 7 \mu\text{m}$. We hypothesize that the discrepancy may be due to the fact that the autocorrelation curves often have ripples and sidelobes that lead to non-Gaussian shapes to the autocorrelation curve. We also note that, for directional DLS-OCT, estimation of v_x is driven by finding a value of v_x that minimizes γ ($1/\gamma$ is the intensity decorrelation time). In contrast, other methods require a more exact estimation of the beam radius because the beam radius serves as a constant of proportionality that relates estimated decorrelation time to scatterer translational speed.

5.3 DLS-OCT total velocimetry using less data and few scan biases

We next investigated the effects of using fewer scan bias velocities ($n_{bias} = 4$) with no repeated data acquisition $n_{data} = 1$. We considered two sets of four scan bias velocities: the

fastest (-13.7 , 13.7 , -6.8 , and 6.8 mm/s; Fig. 7) and the slowest (-3.4 , 3.4 , -1.7 , and 1.7 mm/s; Fig. 8) scan biases. These two sets can be thought of as representing two extremes of partitioning the data—either scanning at velocities close to the flow or far away. Once again, the adaptive hyperprior reduced uncertainty, with Table 1 summarizing the magnitude uncertainty reductions.

In the case of the faster scan velocities, using only an uninformative prior gives reconstructions of relatively poor precision with the uncertainty in measurement being greater than the measurement itself. Using prior information, however, significantly improves precision (Fig. 7 and Table 1). In the case of the slowest scan velocities, Markov chains did not reach steady state and thus did not arrive at a stationary distribution. As a result, we could not reliably compare the improvement of using the hyperprior distribution. In order to reach steady state, we used beam waist as additional prior information in order to more tightly constrain the posterior distribution. Fixing the beam waist parameter value at $5\text{ }\mu\text{m}$ led to Markov chains reaching steady state and stationary posterior distribution estimates (Fig. 8), and use of an adaptive hyperprior also leads to improved estimation precision.

Table 1. Improvements with fewer scan biases

95% CI	fastest biases		slowest biases	
	lateral	axial	lateral	axial
uninformative prior	4510 $\mu\text{m/s}$	63 $\mu\text{m/s}$	2100 $\mu\text{m/s}$	14 $\mu\text{m/s}$
adaptive hyperprior	2270 $\mu\text{m/s}$	39 $\mu\text{m/s}$	1450 $\mu\text{m/s}$	10 $\mu\text{m/s}$
uncertainty reduction	50%	38%	31%	29%

Summary of results from using neighboring information via adaptive hyperpriors for the use of only four scan biases. CI widths are reported as the median across the channel due to a few extreme values. Note that for the slowest scan biases, prior information about the beam waist had to be included.

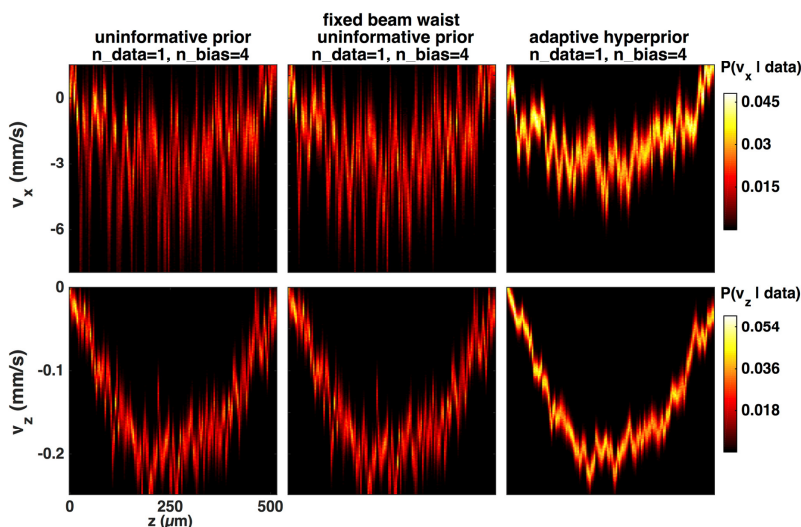


Fig. 7. Bayesian estimates of DLS parameters using 1 sample and the 4 fastest scan bias velocities. The first column consists of results from an uninformative prior; the second, uninformative priors on all parameters except the lateral beam waist; and the third, adaptive hyperprior. Only the second column assumes a fixed beam waist of $5\text{ }\mu\text{m}$.

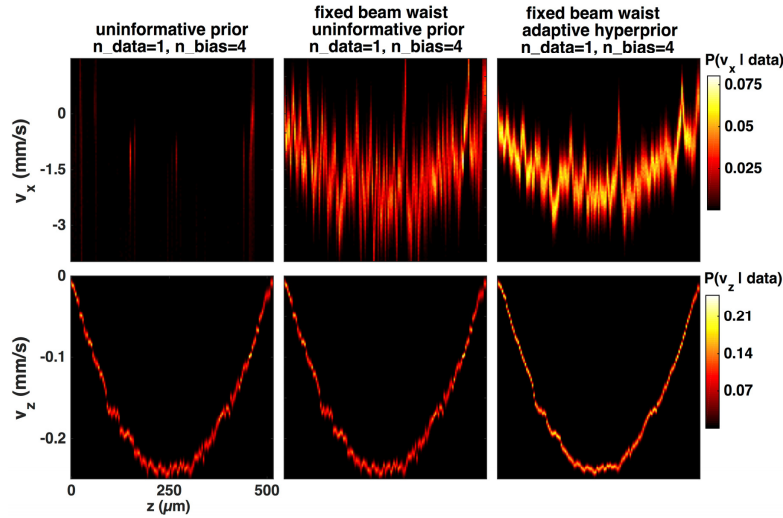


Fig. 8. Bayesian estimates of DLS parameters using 1 sample and the 4 slowest scan bias velocities.

5.4 Axial velocimetry of cardiac motion in *Drosophila melanogaster* (fruit fly) embryos

In order to demonstrate the feasibility of our Bayesian approach in a biomedical context, we applied our method for improving the precision of Doppler velocity estimates of heart wall motion in *Drosophila melanogaster* (fruit fly) pre-pupae. Figure 9(a) shows the M-mode image of a wild-type fruit fly heart. In order to extract the Doppler signal from the heart wall, the walls were segmented using a tracking algorithm. Starting with an initial seeded spatial location manually chosen, the subsequent spatial location and the next time point was chosen based on the maximum intensity across the A-scan with a quadratic penalty for larger distances. To address non-stationarity, a short-time Fourier transform was calculated with a sliding Gaussian window with a full width at half maximum of 360 μs (10 data points). Squaring the magnitude of the windowed Fourier transform followed by inverse Fourier transformation gave an estimate of the autocorrelation signal as per the Wiener-Khinchin theorem. The autocorrelation signal was fit to the DLS model with a single decorrelation parameter capturing the effects of diffusion and translational decorrelation. The window was slid by the half width at half max to avoid overly double-counting data. Although we were interested in the axial velocity v_z , the Bayesian framework we implemented requires analysis of the entirety of Eq. (3), which also incorporates information about lateral velocity. In the context of Bayesian analysis, one can estimate all the parameters required by the model, but integrate over estimates of non-essential parameters. In doing so, one is left with only parameters of interest, in this case, v_z .

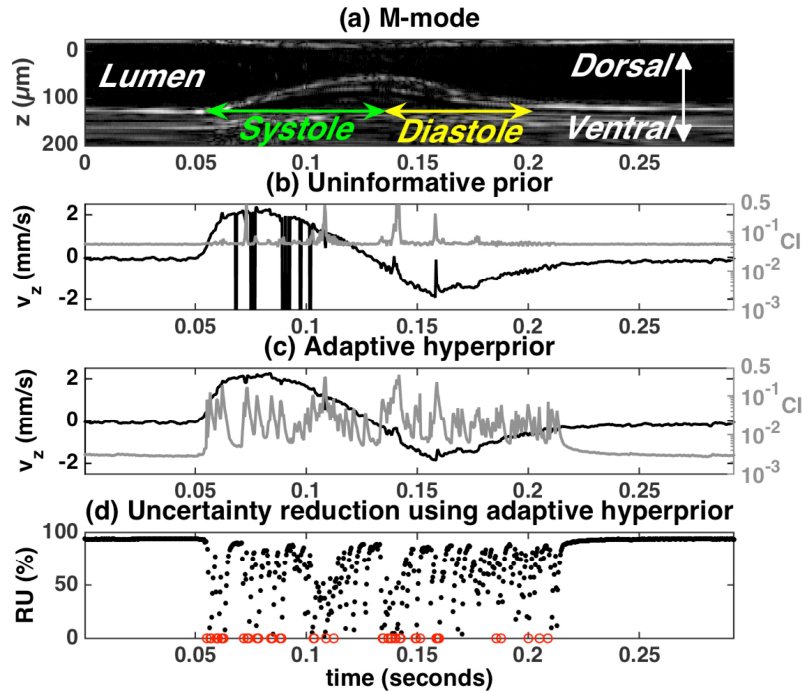


Fig. 9. Bayesian heart wall velocimetry in a *Drosophila melanogaster* pre-pupa. All horizontal axes are identical (time). (a) M-mode image of *D. melanogaster* heart. (b) Heart wall velocity trace (black) and 95% CI for Bayesian analysis using a broad uninformative prior. Note that CI plot is on a log scale. (c) Heart wall velocity trace (black) and 95% CI for Bayesian analysis using an adaptive hyperprior. (d) Reduction in uncertainty when using an adaptive hyperprior compared to a broad uninformative prior. The vertical axis is truncated at $RU = 0$. The uninformative prior occasionally outperformed the adaptive hyperprior (i.e. $RU < 0$). These datapoints are highlighted in red.

The uncertainty reduction was greatest for the adaptive hyperprior when the velocity profile varied slowly, while rapid movements during contraction and relaxation have more variable uncertainty reduction. This pattern of uncertainty reduction may be attributed to the intuitive notion that the slower the parameter variation, the more information the parameter estimates at one location are applicable to those of neighboring locations. Note that there were occasional instances where the uncertainty increased as a result of the incorporation of neighboring information (red points in Fig. 9(d)). This increase may be attributed to the fact that we also adapted the posterior distribution for the data variance parameter (σ^2 in Eq. (6)). Hence, if the neighboring position significantly deviated from the model (a large σ^2), then forcing the data at the current position to have a high variance about the model (even if the data suggest it should not) would allow the fit to vary widely in order to accommodate this larger variance. As such, the posterior distribution in the parameters may widen. The fact that an adaptive hyperprior does not categorically narrow CIs may be viewed as a desirable result. It is desirable because it suggests that if the data in the neighboring position has a high variance about the model (e.g., because of violation of stationarity), then the parameter estimates from that neighbor are less reliable than if the model tightly fit the data. This guardrail against inappropriate narrowing is reflected in a useful quality control rule: in cases where $RU < 0$, CIs generated using an uninformative prior should be used in lieu of CIs generated using the adaptive hyperprior.

6. Conclusion and discussion

We developed a Bayesian framework for OCT velocimetry that reduces uncertainty in velocity profile reconstructions in a calibrated phantom and in an important animal model of human disease. Here, uncertainty is defined as the width of the posterior distribution of the velocity parameter estimate (e.g. $P(v_x | \text{data})$). In this study, we used 95% CIs as a metric for posterior distribution width. Reduction in uncertainty in the Bayesian framework is accomplished through the incorporation of prior information. In particular, we have shown that spatially neighboring locations provide some information about the current location, under the assumption that spatial properties do not vary rapidly between neighbors. The rationale is that if we estimate the parameters in one location, we will have gained information about nearby locations, and our proposed adaptive hyperprior method is one way of parameterizing this information. A major decision was in how strongly we chose to use prior information from adjacent spatial locations to inform our next estimate. In principle, we could have set a more restrictive prior such that the posterior distribution of one location gives the prior distribution for the next location. Using neighboring information in this manner is conceptually similar to averaging and thus would lead to progressively narrower posterior distributions across the channel. Such narrowing is at odds with the fact that flow velocity is expected to slowly vary across an image. Rather, we used an adaptive hyperprior approach. The adaptive hyperprior approach uses the posterior distribution of one location to be the prior distribution of the hyperparameter of the next prior while keeping the variance the same. This approach avoids a gradual narrowing of the posterior distribution as velocity estimates are generated across the image data field of view.

We believe that the primary advantage of the presented approach is the ability to incorporate prior information into a statistical framework for velocity estimation using simpler (e.g. Doppler) or more complex (e.g. directional DLS-OCT) models of the complex-valued OCT signal. The Bayesian approach also yields credible intervals (CI) and posterior distributions ($P(\theta | \text{data})$) that assist in further interpreting velocity estimates. On that point, Kasai estimators do not give confidence or credible intervals. Thus, although Kasai is well-established, it is not straightforward to compare Kasai to methods that yield confidence or credible intervals.

We made a few assumptions in our overall estimation framework. These assumptions reduced the computational burden associated with Bayesian analysis. We believe that these assumptions are reasonable, although future work may focus on more detailed estimation models and OCT signal models. In terms the estimation model, we assumed a homoscedastic Gaussian noise model (i.e. constant noise variance across all autocorrelation lags). A more detailed heteroscedastic noise model would estimate a noise variance for each autocorrelation lag. Second, the deviations of the observed autocorrelation from the proposed model were assumed to be independent and identically distributed Gaussian, even though the uncertainties increase slowly with lag and are potentially correlated due to the fact that each lag calculation uses some overlapping data. To avoid the former issue, we restricted the number of lags to the first 15. Third, all prior distributions used in this study were assumed to be independent of each other. Fourth, in terms of using neighboring locations for prior information, our approach used immediately adjacent spatial locations as sources for priors. Although the immediate neighbor approach is straightforward, we note that there are more general approaches (e.g. Markov random fields [16]) that avoid the directionality (i.e. left-to-right or right-to-left) of the immediate neighbor approach. In terms of the model of the OCT signal, we assumed that the velocity gradient is zero. Future work may focus on using signal models that assume a non-zero gradient as in Weiss, *et al.* [7].

Analysis reported in Fig. 2 revealed a small scanner-induced Doppler shift. This scanner-induced Doppler shift could be incorporated into the velocity estimation process by expanding the v_z term in $\exp(2jkv_z\tau)$ in Eq. (3) to $v_z + mv_{x,scan}$. Here, m is scanner-induced

Doppler shift per unit scan velocity. By analyzing the total Doppler shift as a function of scan bias velocity in Fig. 2, we estimate that the scanner-induced Doppler shift m is $-16.4 \mu\text{m/s}$ per 1 mm/s of scan velocity. However, our expectation was that the scanner-induced Doppler shift does not significantly impact axial velocity measurements. Our expectation was based in the fact that, across all scan biases $v_{x,scan}$, the average scanner-induced Doppler shift is zero. That is, the average value of $(v_z + mv_{x,scan})$ taken across all scan biases is v_z because the set of scan bias values used is symmetric around zero. Our expectation is supported by two observations. First, axial flow velocities at the edges of the tube are zero or near-zero in the various analyses performed. If the scanner-induced Doppler shift imparted a significant velocity bias, this bias would be expected to be present in velocity estimates of stationary scatterers. Second, the ratio of v_z and v_x is consistent with the estimated angular tilt of the capillary tube based on intensity OCT imaging.

As it relates to transverse velocimetry, the salient feature of Eq. (3) is the relationship between the rate of amplitude signal decorrelation and the scan bias velocity. Varying the scan bias velocity modulates the decorrelation rate in a predictable manner and thus provides a baseline set of decorrelation rates. In the presence of scatterer motion parallel (or antiparallel) to the scan bias direction, scatterer velocity along that direction can be inferred from the degree of departure from baseline decorrelation rates. Thus, while the form of the diffusive term is simplified in Eq. (3) with respect to, for example, Lee, *et al.* [4], this simplification enables important new functionality, that is, the isolation of one specific parameter value that is determined on an amplitude (not phase) basis. In principle, if both v_x and v_y were sequentially estimated using scan bias protocols along the x - and y -axes, respectively, D could be inferred from residual amplitude decorrelation not otherwise attributable to v_x and v_y .

Incorporating prior knowledge of v_z (e.g. through phase-sensitive Doppler estimation of v_z) would not change the process for estimating the transverse velocity parallel to the x -axis (v_x). The estimation process is unchanged because the v_x is determined by value of the scan bias ($v_{x,scan}$) that minimizes γ . Here, $1/\gamma$ is the intensity decoration time. Since v_z is invariant with scan velocity, it would not be expected to change the $v_{x,scan}$ at which γ reaches a minimum. In Huang, *et al.* [8], v_x was determined by fitting γ versus $v_{x,scan}$ to a parabola and finding the minimum of that fit curve. Here, the analogous parabolic relationship is present in the numerator of the argument of the Gaussian function in Eq. (3).

Using CI widths as a metric of estimator precision, we note that the Doppler frequency estimates are more precise than decorrelation time-based DLS estimates. The lower precision of DLS estimates is consistent with our subjective experience that decorrelation-based measurements are more susceptible to uncertainty than Doppler frequency measures are. Future studies may focus on understanding why uncertainty is apparently higher with DLS-OCT than with Doppler OCT. Since decorrelation times have a conjugate bandwidth in the Fourier domain, one potential explanation is that first-order moments (Doppler center frequencies) are less susceptible to error compared to second-order moments (Doppler frequency bandwidth). Our quantitative results are consistent with observations previously made by Srinivasan, *et al.* [5].

Acknowledgments

This work was supported by a Basil O'Connor Starter Scholar Research Award Grant No. 5-FY13-211 from the March of Dimes Foundation and NIH 1R01HL118419-01. BKH also was also supported by NIH MSTP TG T32GM07205.